42 Application of Statistical and Neural Classifiers to Recognition of Handwritten Digits

Erkki Oja, Jorma Laaksonen, and Lassi Tuura

Recently, many benchmark and comparison studies have been published on neural and statistical classifiers. One of the most extensive was the Statlog project [4] in which statistical methods, machine learning, and neural networks were compared using a large number of different data sets. As a general conclusion of that study, good statistical classifiers included the k-Nearest Neighbor (k-NN) rule, and good neural network classifiers included Learning Vector Quantization (LVQ) and Radial Basis Function (RBF) classifiers. One of the databases used in that study consisted of handwritten digits. The good performance of nearest neighbor classifiers for handwritten digits was also confirmed by Blue et al.[1], who however only compared the result to RBF and Multi-Layer Perceptron (MLP) neural classifiers. A kernel discriminant analysis (KDA) type method, the Probabilistic Neural Network (PNN), also did very well in that study.

In our study [3], we wanted to compare some classification methods developed within our research group to the standard classifiers. Special interest has been on the subspace methods of classification. The Averaged Learning Subspace Method (ALSM) [5] and some new modifications of it have been used in our experiments and compared against the classification error levels obtained with some other, more commonly used classification methods. As a case study, recognition of handwritten digits was used.

Off-line character recognition is one of the most popular practical applications of pattern recognition. During the last two decades, the research has mostly focused on handwritten characters. This is mainly due to the fact that while recognition of machine printed characters is considered a solved problem, the reliability achieved with handwritten text has not yet reached the level required in practical applications [2]. Handwritten character recognition has become a popular application area for neural network classifiers, which through their adaptive capabilities have often been able to achieve better reliability than classical statistical or knowledge-based structural recognition methods. For these reasons, we have also chosen to use handwritten digit data in our experiments.

The design of the feature extraction stage is an essential step in the development of a complete pattern recognition system. We have chosen to use simple statistical features obtained by applying Karhunen-Loève transformation on the normalized input images. Similar approach has been taken by many research groups in the Second Census OCR Conference [2] and also in the comparison [1]. Some exemplary images of handwritten digits are displayed in the leftmost column of Figure 42. The reconstruction of the original images from the calculated features is seen to get more accurate as the number of feature terms is increased.

Our evaluation study has been carried out by using systematic training set cross-validation in all classifier design. The final performance estimates are based on an independent testing set that has had no role in classifier construction, including the choice of optimal pattern vector dimension. The classification accuracies of

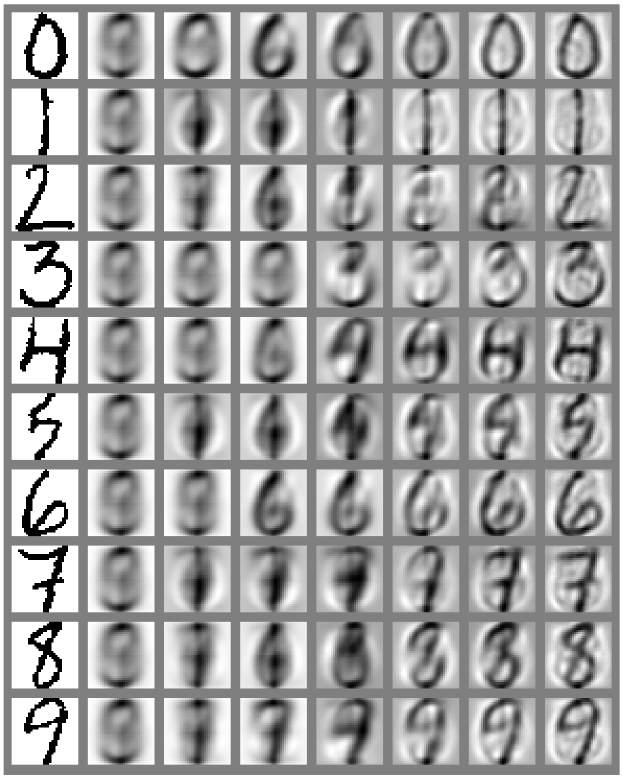

Figure 87: Some examples of handwritten digits and their reconstructions from increasing number of feature terms.

a subset of the classifiers tested are displayed in Table 14. While the case study does give some indication of the relative merits of the tested methods in the two particular applications studied, we want to emphasize that the results obtained by no means provide grounds for any objective ranking of these methods as alternative general classification schemes. The best classifier for a given task can be found by experimenting with different designs and basing the choice on criteria which, in addition to classification error, can include other issues such as computational complexity and feasibility of efficient hardware implementation.

classifier	**error-%**
KDA	3.5
MLP	3.5
LLR	2.8
k-NN	3.8
ALSM	3.2
committee	2.5

Table 14: Some examples of classification error levels obtained in the study.

As the last part of the evaluation, a committee comprising of three different classifiers was formed. By doing a majority vote of the outputs of the three member, the committee classifier was able to lower the error rate to a level below the accuracy of any of the classifiers involved. In the prototype handwritten digit recognition system developed for the current study, the option to reject a difficultly classifiable input image has also been implemented. As a general phenomenon in recognition of handwritten characters, it has been observed [2] that the rejection-error curve is linear in the $\rho \log \epsilon$-plane where ρ and ϵ denote overall rejection and misclassification rates, respectively. This feature has been confirmed by our experiments as seen in Figure 88.

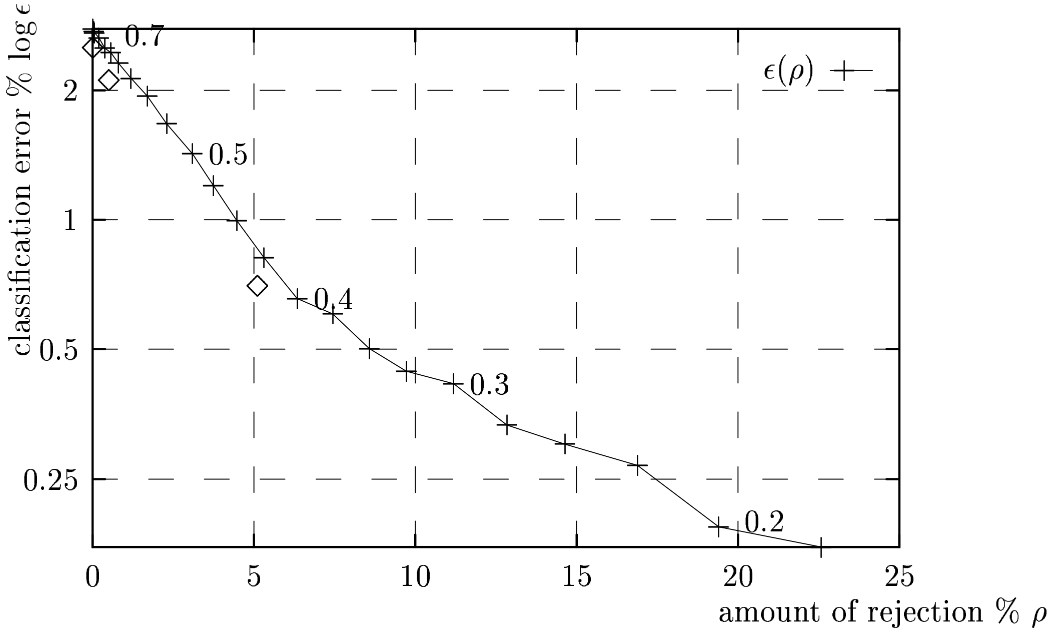

Figure 88: The rejection-error trade-off curve of the LLR classifier.

References

[1] J. L. Blue, G. T. Candela, P. J. Grother, R. Chellappa, and C. L. Wilson. Evaluation of pattern classifiers for fingerprint and OCR applications. *Pattern Recognition*, 27(4):485–501, 1994.

[2] J. Geist, R. A. Wilkinson, S. Janet, P. J. Grother, B. Hammond, N. W. Larsen, R. M. Klear, M. J. Matsko, C. J. C. Burges, R. Creecy, J. J. Hull, T. P. Vogl, and C. L. Wilson. The second census optical character recognition systems conference. Technical Report NISTIR 5452, National Institute of Standards and Technology, Aug. 1992.

[3] D. Michie, D. J. Spiegelhalter, and C. C. Taylor, editors. *Machine learning, neural and statistical classification.* Ellis Horwood Limited, 1994.

[4] E. Oja. *Subspace Methods of Pattern Recognition.* Research Studies Press Ltd., Letchworth, England, 1983.